DOCUMENT RESUME

ED 445 646 HE 033 336

AUTHOR Walsh, Timothy A.
TITLE Identifying Peer Institutions for Graduation Rate Comparisons. AIR 2000 Annual Forum Paper.
PUB DATE 2000-05-00
NOTE 18p.; Paper presented at the Annual Forum of the Association for Institutional Research (40th, Cincinnati, OH, May 21-23, 2000).
PUB TYPE Reports - Research (143) -- Speeches/Meeting Papers (150)
EDRS PRICE MF01/PC01 Plus Postage.
DESCRIPTORS Bachelors Degrees; College Environment; *College Graduates; Colleges; Comparative Analysis; Degrees (Academic); Educational Attainment; *Graduation; Higher Education; *Institutional Characteristics; Institutional Evaluation; *Outcomes of Education; *Peer Institutions; School Demography; Statistical Analysis; Undergraduate Study; Universities
IDENTIFIERS *AIR Forum; *Graduation Rates

ABSTRACT

This study used publicly available information to identify institutional controllable variables that influence graduation rate and to group colleges using these factors. The dependent variable for the correlation analysis was the six-year graduation rate of the fall 1993 freshman cohort (as reported by US News and World Report) at more than 1,400 bachelors-degree granting institutions. Factor analysis identified five components that accounted for 78.8 percent of the graduation rate variance: student academic ability (defined primarily by average combined Scholastic Assessment Test score, freshman retention rate, percent of freshmen from top 25 percent of their high school class, percent of applicants admitted, and number of colleges in geographic area); presence of nontraditional students (percent of part-time students, average age, percent male, percent living in student housing, student/faculty ratio); presence of disadvantaged students (percent minority, percent taking only remedial courses); institution sector (including percent of students on financial aid); and institution mission (liberal arts/comprehensive curriculum focus, doctoral degrees awarded, federal research dollars). Cluster analysis grouped 971 colleges on the five principal components but was found unsuitable for establishing an objective grouping of peer institutions for graduation rate comparisons. (Contains 12 references.) (CH)

IDENTIFYING PEER INSTITUTIONS FOR GRADUATION RATE COMPARISONS

Timothy A. Walsh
Director, Student Information Systems
Temple University
1801 N Broad St
Philadelphia, PA 19122-6096
215/204-5050

IDENTIFYING PEER INSTITUTIONS FOR GRADUATION RATE COMPARISONS

The graduation rate of bachelor's degree granting institutions has drawn considerable interest over the past decade. This study used publicly available data to identify the institutional controllable variables that influence graduation rate and to group colleges using these factors.

Correlation analysis was used to identify the institutional variables associated with the six-year graduation rate of the fall 1993 freshman cohort for over 1400 colleges. Factor analysis identified five principal components that account for 79 percent of the graduation rate variance. The components are student academic ability, presence of non-traditional students, presence of disadvantaged students, institution sector and institution mission.

Cluster analysis was then performed to group 971 colleges based upon their factor scores on the five principle components. Cluster analysis was found to be unsuitable for establishing a completely objective grouping of peer institutions for graduation rate comparisons.

IDENTIFYING PEER INSTITUTIONS FOR GRADUATION RATE COMPARISONS

INTRODUCTION

In recent years there has been considerable interest in college graduation rates. The Integrated Postsecondary Education Data Set (IPEDS) Graduation Rate Survey provides a standardized way of reporting six-year graduation rates. Most of the studies of postsecondary graduation have focused on student characteristics and experiences that influence persistence and graduation. While individual graduation factors are important, many cannot be put into a useful format or are unavailable for institutional comparisons. This study focuses on institution controllable variables associated with graduation. Who should we compare ourselves to? Are there factors that we, as an institution, have control over that will improve our graduation rate? Is there an objective way of identifying our peers for graduation rate comparisons?

Many colleges establish peer groups for bench marking and other comparative activities. Identification of one's peers has been subjective. A number of coalitions have been formed for data sharing (HEDS, AAU, etc.), but each is limited to a select group of institutions. Formation of the coalition may be on factors unrelated to graduation. Some data exchange groups collect specialized information (such as average family income), but again the data are for a limited group of institutions.

OBJECTIVES

This study tried to achieve three objectives:

- Identify institution controllable variables that influence graduation rate.
- Use publicly available data.
- Determine if there is an objective way to group institutions for peer comparison.

PROCEDURE

The dependent variable in this study was the six-year graduation rate of the fall 1993 freshman cohort (reported by *US News and World Report* in fall 1999) at bachelors degree granting institutions. The literature suggests that there are at least twenty institution controllable variables associated with postsecondary graduation. These variables can be categorized as student demographics, student academic ability and institutional characteristics.

Numerous studies have found gender and ethnicity as important student demographic variables (Astin, 1996). Socio-economic status is also important (Lavin, 1997). This study used percent male undergraduates and percent minority (calculated from IPEDS Fall Enrollment data), and percent on financial aid (using College Board data) as measures of these factors. The presence of older students also impacts on a six-year graduation rate (Robb, 1988). The average age of undergraduates (using College Board data) was used to reflect the presence of older students.

Student academic ability is an obvious factor in graduating from college. High school class rank and standardized test scores are often used to make admissions decisions. Two measures of class rank are available as public data - percent of new freshmen in the top ten percent of their high school class and percent of new freshmen in the top 25 percent. Similarly, standardized test scores are reported as the score at the 75th and 25th percentiles for new freshmen. Class rank and standardized test scores are available from College Board data. Standardized test scores are frequently reported as SAT Combined scores. However, a significant number of colleges use the ACT scores. Fortunately, there has been a way to convert ACT scores to SAT scores (Marco and Abdel-Fattah, 1991). Another possible measure of academic ability is the percent of undergraduates taking only remedial courses (calculated from IPEDS Fall Enrollment data).

Institutional characteristics take many forms. Primary is the institution's mission. Sector (public / non-public) and Carnegie classification are good measures of mission (from IPEDS Institutional Characteristics data). Geographic variables such as region of the country, setting (urban/rural) and competition (measured by the number of other bachelors degree institutions in the area) can be obtained from IPEDS Institutional Characteristic data. An institution also has control over its overall size, the size of its undergraduate student body, freshman admissions selectivity, percent of undergraduate classes with under 20 students, student/faculty ratio, percent of the education and general operating budget spent on student services, and percent of undergraduates living in housing. Institutions have less control over the percent of undergraduates attending part-time and the freshman one-year retention rate. College Board data, *US News* data, and IPEDS data were used to populate these variables.

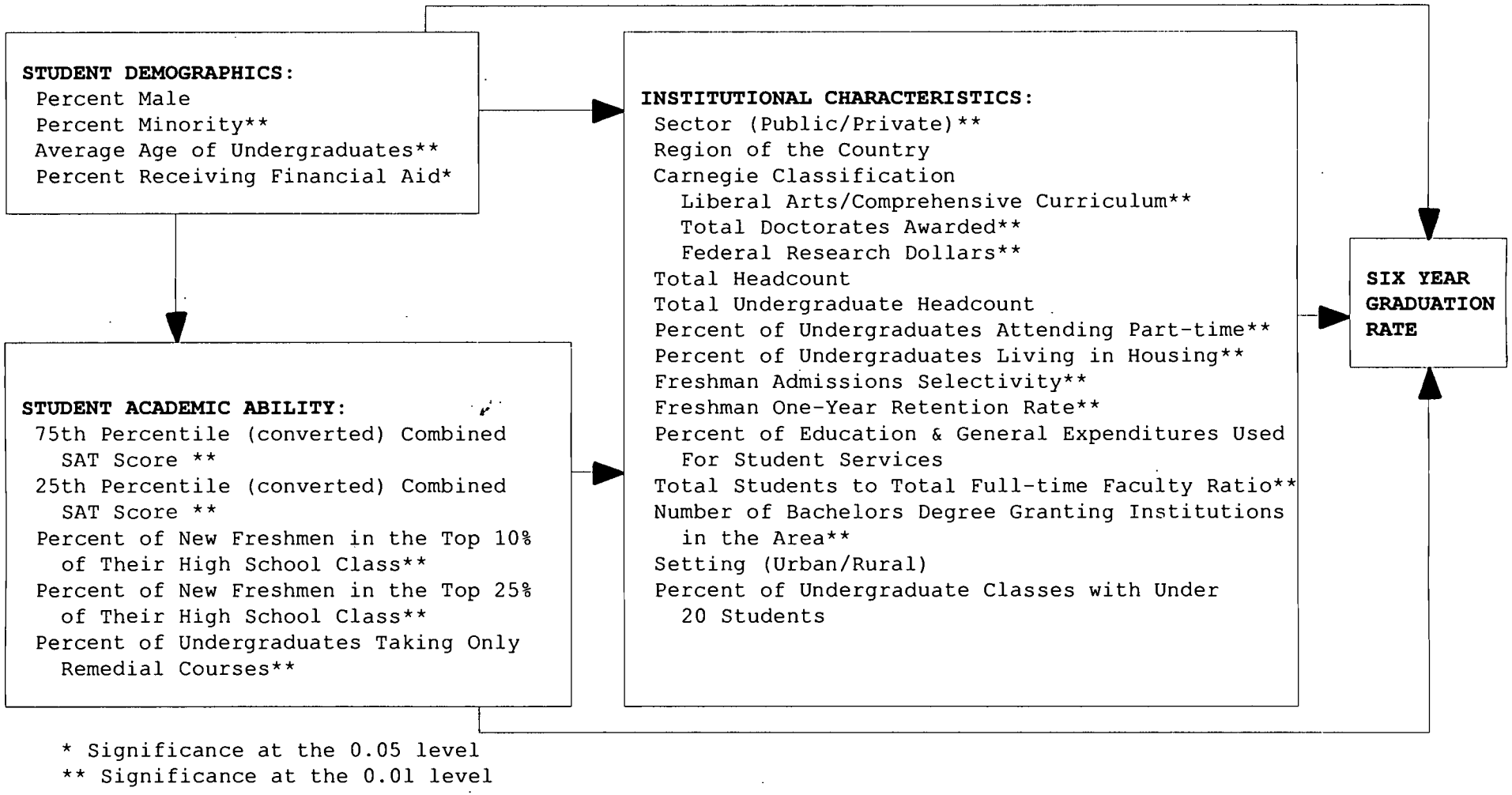
Faculty/ Student Ratio was calculated several different ways using IPEDS Fall Enrollment and Staffing data. Total students to total faculty produces one ratio. Total undergraduates to total faculty produces another, as does full-time undergraduates to full-time faculty. Finally, there is full-time equivalent (FTE) undergraduate students to FTE faculty. All were calculated.

Institutional setting was dichotomist into urban/rural for further analysis. The Carnegie classification code did not lend itself to data analysis, so it was disaggregated into its component parts - liberal arts/comprehensive curriculum focus, doctoral degrees awarded and federal research dollars. Region of the country could not be used in the analysis because it could not be meaningfully recoded for further analysis.

ANALYSIS

Correlation Analysis

Zero order correlations were calculated for 28 variables. Some of the variables were variations of the same measure (SAT scores, high school class rank, student/faculty ratio). The variable with the higher correlation with graduation rate was retained for further analysis. Variables with a .05 level of significance or better were retained in the model. Figure 1 reflects the results of the correlation analysis.



* Significance at the 0.05 level
** Significance at the 0.01 level

FIGURE 1: VARIABLES INFLUENCING GRADUATION RATE

Factor Analysis

Factor analysis was used to identify inter-relationships among the variables. It is useful in providing insight into the underlying structure of the data. The goal was to account for a large part of the variance with as few variables as possible. Discriminate analysis was not used because we did not want to impose a structure on the data through a hypothesis. Instead we let the data define the structure. After data reduction, standardized factor coefficients were used to calculate factor scores for each institution. Stepwise regression was used to assess each factor's contribution to the variance in graduation rate. Table 1 reports the factor coefficients from factor analysis.

Cluster Analysis

Cluster analysis was performed using SPSS. Factor scores define an institution's placement. There are a variety of different methods for doing cluster analysis. The nearest neighbor method was used to group clusters. Cluster analysis does not produce just one, optimal solution. The cluster analysis was repeated several times to bracket-in the optimal number of clusters.

RESULTS

The correlation analysis allowed us to reduce the number of variables used for further analysis. Percent male, total and undergraduate headcount, student services expenditures, setting and

VARIABLE	FACTOR 1	FACTOR 2	FACTOR 3	FACTOR 4	FACTOR 5
25th Percentile SAT Score	0.84690	-0.26256	-0.17215	0.05128	0.14818
Freshman Retention Rate	0.81436	-0.18168	-0.20895	0.01822	0.17952
Freshman Percent in Top 25% of High School Class	0.76341	-0.28008	-0.14062	0.02391	0.20205
% Freshman Applicants Admitted	-0.62659	0.07370	-0.32807	0.14924	-0.03923
Number of Colleges in the Area	0.41572	0.36653	0.30749	0.32326	0.16695
Undergrad Percent Part-Time	-0.18702	0.83352	-0.08815	-0.14633	-0.02128
Average Age of Undergraduates	-0.27530	0.67253	0.07241	-0.20048	-0.00207
UG Student/Full-Time Faculty Ratio	-0.35092	0.61031	-0.15999	-0.21066	-0.12725
Percent of UG in Campus Housing	0.28216	-0.56825	-0.11061	0.56275	-0.15089
Percent Male	-0.07377	-0.55513	-0.09584	-0.29939	0.21777
Percent UG Minority Students	-0.05803	0.03977	0.79708	0.02438	0.14341
Percent of UG Students Taking Only Remedial Coursework	-0.08503	-0.04361	0.61383	-0.11171	-0.14535
Sector (Public=1; Private=2)	0.22791	0.00736	-0.10272	0.83292	-0.10771
Percent UG Receiving Financial Aid	-0.37411	-0.16234	0.05011	0.73515	0.01298
Number of Doctoral Degrees Awarded	0.19584	-0.08900	-0.02923	-0.18584	0.89622
Federal Research Dollars	0.23003	-0.08520	0.02551	-0.15460	0.89131
Comprehensive (1)/Liberal Arts(2)Focus	-0.07085	-0.12714	-0.06106	-0.26697	0.62031
Percent of Variance Explained	**59.5%**	**9.4%**	**4.3%**	**2.8%**	**2.5%**

[ ] - Variables having the most impact on the factor.

TABLE 1: FACTOR COEFFICIENTS FROM FACTOR ANALYSIS

classes with under 20 students were dropped because they failed to reach a .05 level of significance. The 25th percentile SAT scores, the percent of new freshmen in the top 25 percent of their high school class, and the total students to total full-time faculty ratio were kept for further analysis since these variables had a higher correlation with graduation rate than did their counterparts.

Factor analysis identified five principal components associated with graduation rate. These five factors accounted for 78.8 percent of the variance in graduation rate.

- *Academic Ability* - Student academic ability accounted for 59.5% of the variance in graduation rate. This factor was primarily defined by the average combined SAT score, freshman retention rate, percent of freshmen coming from the top 25% of their high school class, percent of applicants admitted, and the number of colleges in the geographic area.

- *Non-Traditional Students* - The proportion of the undergraduate student population that were non-traditional students accounted for an additional 9.4% of the variance. The percent of students attending part-time, average age, percent male, percent living in housing and student / faculty ratio contributed most to this factor.

- *Disadvantaged Students* - The third factor was the proportion of disadvantaged students in the undergraduate population. It accounted for an additional 4.3% of the variance. Percent minority and percent taking only remedial courses were the primary contributors to this factor.

- *Institution Sector* - The institution's sector accounted for another 2.8% to the variance. It was primarily defined by the sector and percent of students on financial aid.

- *Institutional Mission* - The fifth most significant factor was the institutions mission. It accounted for another 2.5% of the variance in graduation rate. Liberal arts/ comprehensive curriculum focus, doctoral degrees awarded and federal research dollars primarily defined this factor.

The factor scores were then calculated for each of the 971 institutions that reported data for all of the relevant variables. Using the disaggregation method of cluster analysis, a variety of groupings were calculated. The results were not very satisfactory. The thirty cluster groupings contained from one to 603 colleges. The clusters can be viewed at www.temple.edu/mdev/SIS on the Internet. There is a certain amount of face validity to the groupings. However, in the large clusters there are some unexpected groupings. To be useful, a college would have to try various numbers of clusters to find a satisfactory comparison group. For the thirty cluster groups that have only one or two colleges, reducing the number of clusters to ten may be needed to identify a meaningful group. Similarly, the large groups may need fifty or sixty clusters to form a meaningful group.

CONCLUSION

Clearly the factor analysis was the most important outcome of the study. Almost eighty percent

of the variance in graduation rate can be explained by the five variables of student academic ability, non-traditional students, disadvantaged students, sector and mission. However, a fifth of the variance is yet to be explained. Perhaps some of the new techniques, such as neural networks, will provide better insight into the factors influencing graduation rate.

There is still considerable within cluster variance in graduation rate in even what appear to be good clusters. Why is the University of Michigan's graduation rate 17 percent higher than the University of Texas - Austin when they appear to have similar undergraduate students?

The cluster analysis was the most disappointing part of the study. While the five factors provided the structure for the clusters, determining the optimal number of clusters is still a judgement call by the investigator. Cluster analysis cannot be used to obtain a strictly objective set of peer institutions.

Because the study was looking at six-year graduation rates, the choice of variables was limited to those available for the fall 1993 cohort. It was disappointing that only about forty percent of the institutions had a full set of data, that proportion should increase in the future as more colleges provide data for public consumption. With the emergence of information on the web and the development of the Common Data Set, more publicly available data is becoming available every year. This year was the first time *US News* included freshman retention rate data. This one piece of data added almost twelve percent to explaining the variance in graduation rate.

This study produced some interesting results that should be investigated further. Why are there significantly higher graduation rates at non-public institutions? Some clusters had within group variance of fifty percent or more. If these are supposed to be similar institutions, why is there still so much variance? Studying the differences in the outliers of such clusters may be worthwhile.

BIBLIOGRAPHY

Anderson, K. L. (1987). Persistence, student background and integration/ commitment: Variation by definition of persistence and institutional type. ERIC Document Reproduction Service No. ED 292 380.

Astin, A. W. Et al (1996). Degree attainment rates at American colleges and universities: Effects of race, gender, and institutional type. ERIC Document Reproduction Service No. ED 400 749.

Atherton, B. (1997). Undergraduate persistence and rates of graduation: A cohort analysis of first-time in college freshmen. ERIC Document Reproduction Service No. ED 417 638.

Carroll, C.D. (1987). College persistence and degree attainment. Washington, DC: National Center for Educational Statistics. (ERIC Document Reproduction Service No. 305 024).

Illinois Board of Higher Education (1999). Persistence, graduation, and time to degree. ERIC Document Reproduction Service No. 431 349.

Kroc, R., Woodward, D., Howard, R., Hull, P. (1995). Predicting graduation rates: A study of land grant, research I and AAU universities. Paper presented at the meeting of The Association for Institutional Research, Boston, MA.

Lavin, D.E. (1997). The social construct of graduation rates: Conceptions of college completion and their socio-political implications. Paper presented at the meeting of the American Educational Research Association, Chicago, IL.

McGrath, A. (Ed.). (1999). America's best colleges. US News & World Report. Washington.

Pascarella, E. and Terenzini, P. (1991). How college affects students. San Francisco: Jossey-Bass.

Porter, O. (1990). Undergraduate completion and persistence at four year colleges and universities. Washington, DC: The National Institute of Independent Colleges and Universities.

Robb, P. and others. (1988). Time to degree in public universities: Factors contributing to the length of time undergraduates take to earn their bachelor's degree, Report 88-12. California State Education Commission. (ERIC Document Reproduction Service No. ED 297 680).

Tinto, V. (1987). Leaving college: Rethinking the causes and cures of student attrition. Chicago: University of Chicago Press.

U.S. Department of Education
Office of Educational Research and Improvement (OERI)
National Library of Education (NLE)
Educational Resources Information Center (ERIC)

NOTICE

Reproduction Basis

☑ This document is covered by a signed "Reproduction Release (Blanket)" form (on file within the ERIC system), encompassing all or classes of documents from its source organization and, therefore, does not require a "Specific Document" Release form.

☐ This document is Federally-funded, or carries its own permission to reproduce, or is otherwise in the public domain and, therefore, may be reproduced by ERIC without a signed Reproduction Release form (either "Specific Document" or "Blanket").